

What is claimed is:

- sub
A/
- 1 1. A method of processing a request in a clustered computer system to
2 organize a plurality of members into a group, the method comprising, in a local
3 member from the plurality of members:
 - 4 (a) locally determining within the local member whether the local
5 member is a subgroup leader for a subgroup with which the local member is
6 associated; and
 - 7 (b) if so, transmitting group data on behalf of the subgroup.
 - 1 2. The method of claim 1, further comprising detecting in the local member
2 whether the group data for the subgroup has already been transmitted by a previous
3 subgroup leader, wherein transmitting the group data by the local member is
4 performed only if the group data has not already been transmitted.
 - 1 3. The method of claim 2, wherein detecting whether the group data for the
2 subgroup has already been transmitted by a previous subgroup leader includes
3 accessing a local indicator that indicates whether the group data for the subgroup has
4 already been transmitted, the method further comprising, in the local member:
 - 5 (a) clearing the local indicator in response to receipt of the request;
6 and
 - 7 (b) setting the local indicator in response to receipt of group data for
8 the subgroup by the local member.
 - 1 4. The method of claim 1, further comprising, in the local member:
 - 2 (a) sending an acknowledgment message during an acknowledgment
3 round;
 - 4 (b) waiting for receipt of an acknowledgment message from each of
5 the plurality of members; and
 - 6 (c) processing the group data after receipt of acknowledgment
7 messages from each of the plurality of members.

1 5. The method of claim 4, further comprising, in the local member:

2 (a) receiving a message identifying a failed member among the
3 plurality of members; and

4 (b) in response to receiving the message, identifying the failed
5 member, locally determining whether the local member is the subgroup leader
6 for the subgroup subsequent to the failure of the failed member, and
7 transmitting group data on behalf of the subgroup using the local member if
8 the group data for the subgroup has not yet been transmitted.

1 6. The method of claim 5, wherein receiving the message identifying the
2 failed member includes receiving a membership change message that is automatically
3 generated responsive to failure of the failed member.

1 7. The method of claim 4, wherein sending the acknowledgment message is
2 performed subsequent to the local member transmitting the group data on behalf of
3 the subgroup if the local member is determined to be the subgroup leader, the method
4 further comprising bypassing the transmitting of the group data by the local member
5 if the local member is determined not to be the subgroup leader.

1 8. The method of claim 1, wherein locally determining within the local
2 member whether the local member is the subgroup leader includes comparing a
3 unique characteristic of the local member with those of the other members associated
4 with the subgroup.

1 9. The method of claim 8, wherein the unique characteristic of the local
2 member includes a member name, and wherein locally determining whether the local
3 member is the subgroup leader includes determining whether the local member is the
4 lowest named member among the members associated with the subgroup.

1 10. The method of claim 1, wherein the request comprises a join request,
2 wherein the plurality of members is partitioned into first and second subgroups, the

3 first group associated with existing members of the group, and the second subgroup
4 associated with new members to be added to the group responsive to the join request.

1 11. The method of claim 1, wherein the request comprises a merge request,
2 wherein the plurality of members is partitioned into a plurality of subgroups, each
3 subgroup associated with a partition, and each partition associated with a subset of the
4 plurality of the members.

00/207" 86E/6960

1 12. A method of processing a request in a clustered computer system to
2 organize a plurality of members into a group, the plurality of members partitioned into
3 a plurality of subgroups, the method comprising:

- 4 (a) transmitting group data on behalf of each subgroup; and
5 (b) locally tracking within each member whether the group data for
6 the subgroup associated with such member has been transmitted.

1 13. The method of claim 12, further comprising

- 2 (a) locally determining within each member whether the member is a
3 subgroup leader for the subgroup with which the member is associated; and
4 (b) if the member is a subgroup leader and the group data for the
5 subgroup associated with the member has not been transmitted, transmitting
6 with the member the group data on behalf of the subgroup.

1 14. The method of claim 13, wherein locally tracking within each member
2 whether the group data for the subgroup associated with such member has been
3 transmitted includes, for each member:

- 4 (a) clearing a local indicator in response to receipt of the request; and
5 (b) setting the local indicator in response to receipt of group data for
6 the subgroup associated with such member.

1 15. A method of processing a request in a clustered computer system to
2 organize a plurality of members into a group, the plurality of members partitioned into
3 a plurality of subgroups, the method comprising, for each subgroup:

4 (a) determining a subgroup leader for such subgroup via a local
5 determination made within each active member associated with such
6 subgroup;

7 (b) determining within the subgroup leader for such subgroup whether
8 group data has already been transmitted on behalf of such subgroup;

9 (c) if group data has not already been transmitted on behalf of such
10 subgroup, transmitting the group data using the subgroup leader;

11 (d) performing an acknowledgment round to detect any failed
12 members; and

13 (e) repeating the determination of the subgroup leader for such
14 subgroup, the determination of whether group data has already been
15 transmitted, the transmission of the group data if no such group data has been
16 transmitted, and the performance of the acknowledgment round until no failed
17 members are detected in response to a last performed acknowledgment round.

1 16. An apparatus, comprising:

2 (a) a memory; and

3 (b) a program resident in the memory, the program configured to
4 process a request in a clustered computer system to organize a plurality of
5 members into a group by locally determining for a local member among the
6 plurality of members whether the local member is a subgroup leader for a
7 subgroup with which the local member is associated, and if so, transmitting
8 group data on behalf of the subgroup.

1 17. The apparatus of claim 16, wherein the program is further configured to
2 detect whether the group data for the subgroup has already been transmitted by a
3 previous subgroup leader, and wherein the program is configured to transmit the
4 group data only if the group data has not already been transmitted.

1 18. The apparatus of claim 17, wherein the program is configured to detect
2 whether the group data for the subgroup has already been transmitted by a previous
3 subgroup leader by accessing a local indicator that indicates whether the group data
4 for the subgroup has already been transmitted, and wherein the program is further
5 configured to clearing the local indicator in response to receipt of the request and set
6 the local indicator in response to receipt of group data for the subgroup.

1 19. The apparatus of claim 16, wherein the program is further configured to
2 send an acknowledgment message during an acknowledgment round, wait for receipt
3 of an acknowledgment message from each of the plurality of members, and process
4 the group data after receipt of acknowledgment messages from each of the plurality of
5 members.

1 20. The apparatus of claim 19, wherein the program is further configured to
2 receive a message identifying a failed member among the plurality of members, and
3 in response to receiving the message, identify the failed member, locally determine
4 whether the local member is the subgroup leader for the subgroup subsequent to the

004207 86E26960

21

5 failure of the failed member, and transmit group data on behalf of the subgroup if the
6 group data for the subgroup has not yet been transmitted.

1 21. The apparatus of claim 16, wherein the program is configured to locally
2 determine whether the local member is the subgroup leader by determining whether
3 the local member is a lowest named member among the members associated with the
4 subgroup.

00/201" 86E/6960

22. A clustered computer system, comprising:

- (a) a plurality of nodes coupled to one another over a network;
- (b) a plurality of member jobs defining a group and configured to be executed by at least one of the plurality of nodes; and
- (c) a program configured to be executed by at least one of the plurality of nodes to process a request received by a member job from the plurality of member jobs to add another member job to the group by locally determining for the member job whether the member job is a subgroup leader for a subgroup with which the member job is associated, and if so, transmitting group data on behalf of the subgroup.

1 23. A program product, comprising:

2 (a) a program configured to process a request in a clustered computer
3 system to organize a plurality of members into a group by locally determining
4 for a local member among the plurality of members whether the local member
5 is a subgroup leader for a subgroup with which the local member is associated,
6 and if so, transmitting group data on behalf of the subgroup; and

7 (b) a signal bearing medium bearing the program.

1 24. The program product of claim 23, wherein the signal bearing medium
2 includes at least one of a recordable medium and a transmission medium.